VECTORS FOR TRANSFERRING NUCLEIC ACIDS, COMPOSITIONS CONTAINING THEM AND THEIR USES

The present invention relates to new vectors and to their use for transferring nucleic acids. More particularly, the invention relates to new vectors capable of directing nucleic acids towards cells or specific cellular compartments.

The transfer of nucleic acids is a technique which forms the basis of all the major applications of biotechnology, and increasing the efficiency of transferring nucleic acids constitutes a very important challenge for the development of these applications. The efficiency of transferring nucleic acids depends on numerous factors, among which are the capacity of the nucleic acids to cross the plasma membrane and their capacity to be transported within the cell up to the nucleus.

One of the major obstacles to the efficiency of transferring nucleic acids stems from the fact that the genetic information is often poorly or not directed towards the target organ for which it is intended. Moreover, once the nucleic acid has penetrated into the target cell, it must still be directed towards the nucleus so as to be expressed there. Furthermore, in the case of the transfer of nucleic acids into

differentiated or quiescent cells, the nucleus is delimited by a nuclear envelope which constitutes an additional barrier to the passage of these nucleic acids.

The recombinant viruses used as vectors possess sophisticated and efficient mechanisms for guiding the nucleic acids up to the nucleus. However, viral vectors have certain disadvantages inherent to their viral nature, which unfortunately cannot be completed excluded. Another strategy consists either in transfecting naked DNA, or in using nonviral agents capable of promoting the transfer of DNA into eukaryotic cells. However, the nonviral vectors do not possess subcellular or nuclear targeting signals. Thus the passage of naked DNA, or in combination with its nonviral agent, from the cytoplasm to the nucleus is, for example, a step which has a very low efficiency (Zabner et al., 1995).

Various attempts to attach targeting signals have thus been made. In particular, peptide fragments for targeting have been covalently attached to oligonucleotides in an antisense-oligonucleotide-targeting strategy [Eritja et al., *Synthesis of defined peptide-oligonucleotide hybrids containing a nuclear transport signal sequence*, Tetrahedron, Vol. 47, No. 24, pp. 4113-4120, 1991]. The complexes thus formed

are good candidates as potential inhibitors of expression of endogenous genes.

Transfection vectors comprising a synthetic polypeptide coupled by electrostatic interactions to a DNA sequence have also been described in patent application WO 95/31557, the said polypeptide consisting of a polymeric chain of basic amino acids, of an NLS peptide and of a hinge region which connects the NLS peptide to the polymeric chain and makes it possible to avoid steric interactions. However, this type of construct poses a problem of stability because the interactions called into play between the DNA and the targeting signal are of an electrostatic nature.

Moreover, nucleic acid-specific targeting peptide chimeras exist which are described in patent application WO 95/34664, the binding between the two being of a chemical nature. However, this method involves in particular enzymatic steps which are difficult to control and which do not make it possible to produce large quantities of nucleic acids.

Finally, it has been shown that it is possible to attach an NLS ("Nuclear Localization Signal") sequence to a plasmid DNA via a cyclopropapyrroloindole (Nature Biotechnology, Volume 16, pp. 80-85, January 1998). However, a complete inhibition of the transcription of the gene of

interest because of the random attachment of several hundreds of NLS sequences onto the plasmid was observed. One solution proposed by the authors to remedy this consists in linking the NLS sequences onto linear fragments of DNA and then in coupling these modified fragments with other nonmodified fragments. However, this technique, like the preceding one, has the disadvantage of involving at least one enzymatic step.

Thus, all the methods proposed up until now do not make it possible to solve satisfactorily the difficulties linked to the targeting of double-stranded DNAs.

The present invention provides an advantageous solution to these problems. More particularly, the present invention uses oligonucleotides which are conjugated with targeting signals and which are capable of forming triple helices with one or more specific sequences present on a double-stranded DNA molecule.

Such a vector has the advantage of being able to direct a double-stranded DNA towards cells or specific cellular compartments by means of the targeting signal, without genetic expression being inhibited. The applicant has indeed shown that, by virtue of the formation of stable site-specific triple

helices, it is now possible to link a targeting signal to a double-stranded DNA in a site-specific manner. Consequently, it is possible to attach the targeting signal outside the expression cassette for the gene to be transferred. The applicant has thus shown that genetic expression in the cell is not inhibited in spite of the chemical modification of the DNA. Furthermore, the presence of a triple helix as a means for binding the targeting signal to the DNA is particularly advantageous because it makes it possible to preserve a DNA size which is appropriate for the transfection.

The vector obtained has, furthermore, the advantage of incorporating targeting signals which are very stably linked to the double-stranded DNA, in particular when the oligonucleotide capable of forming the triple helix is modified by the presence of an alkylating agent.

Another advantage of the invention is to make it possible to couple the DNA to be transferred to targeting signals whose number and nature are both controlled. Indeed, it is possible to control the number of targeting signals linked to each double-stranded DNA molecule by introducing a suitable number of specific sequences appropriate for the formation of triple helices into the said double-stranded DNA

molecule. Likewise, it is possible to introduce into the same double-stranded DNA molecule several oligonucleotides linked to different targeting signals (intracellular and/or extracellular), and in this case it is also possible to determine beforehand the respective proportions thereof. In addition, these various targeting signals may be attached to the double-stranded DNA molecule in a manner stable to a greater or lesser degree depending on whether the triple helix is formed with or without a covalent bond (that is to say with or without the use of an alkylating agent).

Finally, the functionalized triple helix obtained results only from chemical conversion steps and may therefore be obtained in a simple manner, reproducibly and in very large, in particular industrial, quantities.

A first subject of the invention therefore relates to a vector useful in transfection, which is capable of targeting a cell and/or a specific cellular compartment. More particularly, the vector according to the invention comprises a double-stranded DNA molecule and at least one oligonucleotide which is coupled to a targeting signal and which is capable of forming, by hybridization, a triple helix with a specific sequence present on the said double-stranded DNA molecule.

For the purposes of the invention, "double-stranded DNA" is understood to mean a double-stranded deoxyribonucleic acid which may be of human, animal, plant, bacterial or viral origin, and the like. It may be obtained by any technique known to persons skilled in the art, and in particular by screening a library, by chemical or enzymatic synthesis of sequences obtained by screening libraries. It may be chemically or enzymatically modified.

This double-stranded DNA may be in linear or circular form. In the latter case, the double-stranded DNA may be in a supercoiled or relaxed state. Preferably, the DNA molecule is of circular form and in a supercoiled conformation.

The double-stranded DNA may also carry a replication origin which is functional or otherwise in the target cell, one or more marker genes, sequences for regulating transcription or replication, genes of therapeutic interest, anti-sense sequences which are modified or otherwise, regions for binding to other cellular components, and the like. Preferably, the double-stranded DNA comprises an expression cassette consisting of one or more genes of interest under the control of one or more promoters and of a transcriptional terminator which are active in the target cells.

For the purposes of the invention, "expression cassette for a gene of interest" is understood to mean a DNA fragment which may be inserted into a vector at specific restriction sites. The DNA fragment comprises a nucleic acid sequence encoding an RNA or a polypeptide of interest and comprises, in addition, the sequences necessary for the expression (enhancer(s), promoter(s), polyadenylation sequences and the like) of the said sequence. The cassette and the restriction sites are designed to ensure insertion of the expression cassette into an open reading frame appropriate for transcription and translation.

It is generally a plasmid or an episome carrying one or more genes of therapeutic interest. By way of example, there may cited the plasmids described in Patent Applications WO 96/26270 and WO 97/10343 which are incorporated into the present by reference.

For the purposes of the invention, gene of therapeutic interest is understood to mean in particular any gene encoding a protein product having a therapeutic effect. The therapeutic product thus encoded may be in particular a protein or peptide. This protein product may be homologous in relation to the target cell (that is to say a product which is normally expressed in the target cell when the latter has no pathological condition). In this case, the expression

of a protein makes it possible, for example, to palliate an insufficient expression in the cell or the expression of a protein which is inactive or weakly active because of a modification, or to overexpress the said protein. The gene of therapeutic interest may also encode a mutant of a cellular protein, having increased stability, a modified activity, and the like. The protein product may also be heterologous in relation to the target cell. In this case, an expressed protein may, for example, supplement or provide an activity which is deficient in the cell, allowing it to combat a pathological condition, or to stimulate an immune response.

Among the therapeutic products for the purposes of the present invention, there may be mentioned more particularly enzymes, blood derivatives, hormones, lymphokines [interleukins, interferons, TNF, and the like (FR 92/03120)], growth factors, neurotransmitters or their precursors or synthesis enzymes, trophic factors [BDNF, CNTF, NGF, IGF, GMF, αFGF, βFGF, NT3, NT5, HARP/pleiotrophin, and the like, dystrophin or a minidystrophin (FR 91/11947)], the CFTR protein associated with cystic fibrosis, tumour suppressor genes [p53, Rb, Rap1A, DCC, k-rev, and the like (FR 93/04745)], the genes encoding factors involved in coagulation [factors VII, VIII, IX], the

genes involved in DNA repair, suicide genes [thymidine kinase, cytosine deaminase], the genes for haemoglobin or other protein carriers, the genes corresponding to the proteins involved in the metabolism of lipids, of the apolipoprotein type chosen from apolipoproteins A-1, A-II, A-IV, B, C-I, C-II, C-III, D, E, F, G, H, J and apo(a), metabolic enzymes such as, for example, lipoprotein lipase, hepatic lipase, lecithin cholesterol acyl transferase, 7-alpha cholesterol hydroxylase, phosphatidic acid phosphatase, or lipid transfer proteins such as cholesterol ester transfer protein and phospholipid transfer protein, an HDL-binding protein or a receptor chosen, for example, from the LDL receptors, the remnant chylomicron receptors and the scavenger receptors, and the like.

The DNA of therapeutic interest may also be a gene or an anti-sense sequence, whose expression in the target cell makes it possible to control the expression of genes or the transcription of cellular mRNAs. Such sequences can, for example, be transcribed in the target cell into RNAs which are complementary to cellular mRNAs and thus block their transcription to protein, according to the technique described in Patent EP 140 308. The genes of therapeutic interest also comprise the sequences encoding ribozymes, which are capable of selectively destroying target RNAs (EP 321

201), or the sequences encoding single-chain intracellular antibodies such as, for example, ScFv.

As indicated above, the deoxyribonucleic acid may also comprise one or more genes encoding an antigenic peptide, which is capable of generating an immune response in humans or in animals. In this specific embodiment, the invention therefore allows the production of vaccines or the carrying out of immunotherapeutic treatments applied to humans or to animals, in particular against microorganisms, viruses or cancers. They may be in particular antigenic peptides specific for the Epstein-Barr virus, the HIV virus, the hepatitis B virus (EP 185 573), the pseudo-rabies virus, the syncitia forming virus, the influenza virus, the cytomegalovirus (CMV), other viruses, or specific for tumours (EP 259 212).

Preferably, the deoxyribonucleic acid also comprises sequences allowing the expression of the gene of therapeutic interest and/or the gene encoding the antigenic peptide in the desired cell or organ. They may be sequences which are naturally responsible for the expression of the gene considered when the sequences are capable of functioning in the infected cell. They may also be sequences of plant origin (responsible for the expression of other proteins, or even synthetic). In particular, they may be promoter

sequences of eukaryotic or viral genes. For example, they may be promoter sequences derived from the genome of the cell which it is desired to infect. Likewise, they may be promoter sequences derived from the genome of the virus. In this regard, there may be mentioned, for example, the promoters of the E1A, MLP, CNV and RSV genes, and the like. In addition, these expression sequences may be modified by the addition of activating or regulatory sequences, and the like. The promoter may also be inducible or repressible.

A triple helix corresponds to the attachment of an oligonucleotide, modified or otherwise, to the double-stranded DNA by so-called "Hoogsteen" hydrogen bonds between the bases of the third strand and those of the region forming the double helix. These pairings occur in the large spiral of the double helix and are specific for the sequence considered [Frank-Kamenetski, M.D., *Triplex DNA Structures*, Ann. Rev. Biochem., 1995, 64, pp 65-95]. The specific sequence in the form of a double helix may be in particular a homopurine-homopyrimidine sequence. Two categories of triple helices can be distinguished according to the nature of the bases of the third strand [Sun, J. and C. Hélène, *Oligonucleotide-directed triple-helix formation*, Curr. Opin. Struct. Biol., 1993, 3, pp. 345-356]: the purine bases make it possible to obtain C-G*G and T-A*A

pairings, and the pyrimidine bases make it possible to obtain C-G*G+ and T-A*T pairings (the symbol * corresponds to a pairing with the third strand).

These structures have been characterized from the physicochemical point of view by means of numerous NMR (Nuclear Magnetic Resonance), hybridization temperature or nuclease protection studies, which makes it possible to define their properties and the conditions for their stability. For the triple helices with the third purine strand, this strand is antiparallel relative to the purine strand of the DNA and the formation of the triple helix highly depends on the concentration of divalent ions: ions such as Mg^{2+} stabilize the structure formed with the third strand. For the triple helices with a third homopyrimidine strand, the latter is parallel relative to the purine strand, and the formation of the triple helix is dependent on the pH: an acidic pH of less than six allows protonation of the cytosines and the formation of an additional hydrogen bond stabilizing the triplet C-G*C+. "Mixed" triple helices also exist for which the third strand carries purine and pyrimidine bases. In this case, the orientation of this third strand depends on the base sequence of the homopurine region.

The oligonucleotides used in the present invention are oligonucleotides which hybridize directly

with the double-stranded DNA. These oligonucleotides may contain the following bases:

- thymidine (T), which is capable of forming triplets with the doublets A.T of the double-stranded DNA (Rajagopal et al., Biochem 28 (1989) 7859);

- adenine (A), which is capable of forming triplets with the doublets A.T of the double-stranded DNA;

- guanine (G), which is capable of forming triplets with the doublets G.C of the double-stranded DNA;

- protonated cytosine (C^+), which is capable of forming triplets with the doublets G.C of the double-stranded DNA (Rajagopal et al., cited above);

- uracil (U), which is capable of forming triplets with the base pairs A.U or A.T.

To allow the formation of a triple helix by hybridization, it is important that the oligonucleotide and the specific sequence present on the DNA are complementary. In this regard, in order to obtain the best attachment and the best selectivity, an oligonucleotide and a specific sequence which are perfectly complementary are used for the vector according to the invention. This may be in particular an oligonucleotide poly-CTT and a specific sequence poly-GAA. By way of example, there may be mentioned the

oligonucleotide having the sequence: 5'-GAGGCTTCTTCTTCTTCTTCTTCTT-3' (GAGG$(CTT)_7$, SEQ ID No.1), in which the bases GAGG do not form triple helices but make it possible to separate the oligonucleotide from the coupling arm. The sequence $(CTT)_7$ (SEQ ID No.2) may also be mentioned. These oligonucleotides are capable of forming a triple helix with a specific sequence comprising complementary units (GAA). These may be in particular a region comprising 7, 14 or 17 GAA units. Another specific sequence of interest is the sequence: 5'-AAGGGAGGGAGGAGAGGAA-3' (SEQ ID No.3). This sequence forms a triple helix with the oligonucleotides: 5'-AAGGAGAGGAGGGAGGGAA-3' (SEQ ID No.4) or 5'-TTGGTGTGGTGGGTGGGTT-3' (SEQ ID No.5).

In this case, the oligonucleotide binds in an orientation which is antiparallel to the polypurine strand. These triple helices are unstable in the presence of Mg^{2+} as is mentioned above (Vasquez et al., Biochemistry, 1995, 34, 7243-7251; Beal and Dervan, Science, 1991, 251, 1360-1363).

The specific sequence may be a sequence which is naturally present on the double-stranded DNA, or a synthetic sequence or a sequence of natural origin artificially introduced into it. It is particularly advantageous to use an oligonucleotide capable of forming a triple helix with a sequence which is

naturally present on the double-stranded DNA. Indeed, this advantageously makes it possible to obtain the vectors according to the invention with unmodified plasmids, in particular commercial plasmids of the pUC, pBR322 and pSV type, and the like. Among the natural homopurine-homopyrimidine sequences present in the double-stranded DNA, there may be mentioned a sequence comprising all or part of the sequence 5'-CTTCCCGAAGGGAGAAAGG-3' (SEQ ID No.6) present in the replication origin ColE1 of E. coli. In this case, the oligonucleotide forming the triple helix has the sequence: 5'-GAAGGGTTCTTCCCTCTTTCC-3' (SEQ ID No.7) and binds alternately to the two strands of the double helix, as described by Beal and Dervan (J. Am. Chem. Soc. 1992, 114, 4976-4982) and Jayasena and Johnston (Nucleic Acids Res. 1992, 20, 5279-5288). There may also be mentioned the sequence 5'-GAAAAAGGAAGAG-3' (SEQ ID No.8) of the β-lactamase gene of the plasmid pBR322 (Duval-Valentin et al., Proc. Natl. Acad. Sci. USA, 1922, 89, 504-508). Another sequence is AAGAAAAAAAAGAA (SEQ ID No. 9) present in the replication origin γ of the plasmids with a conditional replication origin such as pCOR.

Although perfectly complementary sequences are preferred, it is understood, however, that certain mismatches may be tolerated between the sequence of the

oligonucleotide and the sequence present on the DNA, as long as they do not lead to an excessive loss of affinity. There may be mentioned the sequence 5'-AAAAAAGGGAATAAGGG-3' (SEQ ID No. 10) present in the E. coli β-lactamase gene. In this case, the thymine interrupting the polypurine sequence may be recognized by a guanine of the third strand, thus forming a triplet ATG which is stable when it is surrounded by two TAT triplets (Kiessling et al., Biochemistry, 1992, 31, 2829-2834).

The oligonucleotide used may be natural (composed of natural bases which are unmodified or chemically modified). In particular, the oligonucleotide may advantageously exhibit certain chemical modifications which make it possible to increase its resistance or its protection in relation to nucleases, or its affinity in relation to the specific sequence or which also make it possible to provide other additional properties (J. Goodchild, *Conjugates of Oligonucleotides and Modified Oligonucleotides: A Review of their Synthesis and Properties*, Bioconjugate Chemistry, Vol. 1 No. 3, 1990, pp. 165-187).

According to the present invention, oligonucleotide is also understood to mean any succession of nucleosides having undergone a

modification of the backbone. Among the possible modifications, there may be mentioned oligonucleotide phosphorothioates which are capable of forming triple helices with DNA (Xodo et al., Nucleic Acids Res., 1994, 22, 3322-3330), likewise the oligonucleotides possessing formacetal or methylphosphonate backbones (Matteucci et al., J. Am. Chem. Soc., 1991, 113, 7766-7768). It is also possible to use the oligonucleotides synthesized with α-anomers of nucleotides, which also form triple helices with DNA (Le Doan et al., Nucleic Acids Res., 1987, 15, 7749-7760). Another modification of the backbone is the phosphoramidate bond. There may be mentioned, for example, the internucleotide bond N3'-P5' phosphoramidate described by Gryaznov and Chen, which gives oligonucleotides forming particularly stable triple helices with DNA (J. Am. Chem. Soc., 1994, 116, 3143-3144). Among the other modifications of the backbone, it is also possible to mention the use of ribonucleotides, of 2'-O-methylribose, of phosphotriester, and the like (Sun and Hélène, Curr. Opinion Struct. Biol., 116, 3143-3144). The phosphorus-containing backbone can finally be replaced with a polyamide backbone as in the PNAs (Peptide Nucleic Acid), which can also form triple helices (Nielson et al., Science, 1991, 254, 1497-1500; Kim et al., J. Am Chem. Soc., 1993, 115, 6477-6481) or with a guanidine-

based backbone, as in the DNG (deoxyribonucleic guanidine, Proc. Natl. Acad. Sci. USA, 1995, 92, 6097-6101), polycationic analogues of DNA, which also form triple helices.

The thymine of the third strand may also be replaced with a 5-bromouracil, which increases the affinity of the oligonucleotide for DNA (Povsic and Dervan, L. Am. Chem. Soc., 1989, 111, 3059-3061). The third strand may also contain non natural bases among which there may be mentioned 7-deaza-2'-deoxyxanthosine (Milligan et al., Nucleic Acids Res., 1993, 21, 327-333), 1-(2-deoxy-β-D-ribofuranosyl)-3-methyl-5-amino-1*H*-pyrazolo[4,3-*d*]pyrimidine-7-one (Koh and Dervan, J. Am. Chem. Soc., 1992, 114, 1470-1478), 8-oxoadenine, 2-aminopurine, 2'-O-methylpseudoisocytidine, or any other modification known to persons skilled in the art (for a review see Sun and Hélène, Curr. Opinion Struct. Biol., 1993, 3, 345-356).

The object of another type of modification of the oligonucleotide is more particularly to improve the interaction and/or the affinity between the oligonucleotide and the specific sequence. In particular, a completely advantageous modification consists in coupling an alkylating agent to the oligonucleotide. The binding may take place either chemically or photochemically by means of a

photoreactive functional group. Advantageous alkylating agents are in particular photoactivable alkylating agents, for example psoralens. Under the action of light, they form covalent bonds at the level of the pyrimidine bases of the DNA. When these molecules are intercalated at the level of the 5'-ApT-3' or 5'-TpA-3' sequences in a double-stranded DNA fragment, they form bonds with both strands. This light-induced binding reaction can occur at a specific site of the plasmid.

As has been highlighted previously, one advantage of the present invention is therefore the possibility of forming very stable and site-specific triple helices between the oligonucleotide and a specific sequence of the double-stranded DNA by means of a covalent bond formed via an alkylating agent.

The length of the oligonucleotide used in the method of the invention is at least 3 bases, and preferably between 5 and 30 bases. An oligonucleotide having a length of between 10 and 30 bases is advantageously used. The length may of course be adapted on a case-by-case basis by persons skilled in the art according to the desired selectivity and stability of the interaction.

The oligonucleotides according to the invention may be synthesized by any known technique. In particular, they may be prepared by means of nucleic

acid synthesizers. Any other method known to persons skilled in the art may quite obviously be used.

For the purposes of the invention, "targeting signal" is understood to mean targeting molecules of a varied nature. It represents in most cases peptides known for targeting. They may be used to interact with a component of the extracellular matrix, a plasma membrane receptor, in order to target an intracellular compartment or in order to improve the intracellular flow of DNA, during the nonviral transfer of genes in gene therapy.

These targeting signals may comprise, for example, growth factors (EGF, PDGF, TGFβ, NGF, IGF, I, FGF), cytokines (IL-1, IL-2, TNF, Interferon, CSF), hormones (insulin, growth hormone, prolactin, glucagon, thyroid hormone, steroid hormones), sugars which recognize lectins, immunoglobulins, ScFv's, transferrin, lipoproteins, vitamins such as vitamin B12, peptide or neuropeptide hormones (tachykinins, neurotensin, VIP, endothelin, CGRP, CCK, and the like), or any unit recognized by the integrins, for example the peptide RGD, or by other extrinsic proteins of the cell membrane.

It is also possible to use whole proteins, or peptide sequences derived from these proteins, or alternatively peptides which bind to their receptor and

which are obtained by the "phage display" technique or by combinatory synthesis.

Intercellular targeting signals may also be envisaged. Many nuclear homing sequences (NLS) of varied amino acid compositions have been identified and make it possible to get different proteins involved in the nuclear transport of proteins or of nucleic acids. Among these are in particular short sequences (the NLS of the SV40 T antigen (PKKKRKV, SEQ ID No. 11) is one example), bipartite sequences (the NLS of nucleoplasmin which contains two essential domains for nuclear transport: KRPAATKKAGQAKKKKLDK, SEQ ID No. 12), or the M9 sequence (NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY, SEQ ID No. 13) of the hnRNPA1 protein. The proteins carrying these NLS sequences bind to specific receptors, such as the receptors of the importin or karyopherin family, for example. The role of these sequences is to direct the DNA inside the nucleus where it is then immediately available for the transcription machinery, and may be expressed.

The "mixed" targeting signals, that is to say which can serve both for intracellular and extracellular targeting, also come within the scope of the present invention. It is possible to mention, for example, the sugars which target lectins which are present on the cell membrane but also at the level of

the nuclear pores. Targeting by these sugars therefore also relates both to extracellular targeting and nuclear import.

Other signals are involved in mitochondrial targeting (for example the N-terminal part of rat ornithine transcarbamylase (OTC) allows the targeting of mitochondria) or in homing onto the endoplasmic reticulum. Finally, some signals allow nuclear retention or retention at the level of the endoplasmic reticulum (such as the sequence KDEL).

Advantageously, the targeting signals according to the invention make it possible to direct the double-stranded DNA specifically towards certain cells or certain cellular compartments. By way of example, the targeting signals according to the invention can target receptors or ligands at the cell surface, in particular the receptors for insulin, for transferin, for folic acid, or any other growth factor, cytokines, or vitamins, or particular polysaccharides at the surface of the cell or on the neighbouring extracellular matrix.

The synthesis of oligonucleotide-targeting signal chimaeras occurs on a solid phase or in solution and takes into account the very different stability properties of the oligonucleotides and targeting signals (Erijita, R. et al., *Synthesis of defined*

peptide-oligonucleotide hybrids containing a nuclear transport signal sequence, Tetrahedron, 1991, 47(24), pp. 4113-4120]. In solution, it is possible to envisage couplings in one step: the targeting signal may, for example, be synthesized with a group carrying disulphide, maleimide, amine, carboxyl, ester, epoxide, cyanogen bromide or aldehyde functional groups, and may be coupled to an oligonucleotide modified by a thiol, amine or carboxyl terminal group at the 3' or 5' position. These couplings are made by establishing disulphide, thioether, ester, amide or amine bonds between the oligonucleotide and the targeting signal. Any other method known to persons skilled in the art may be used, such as bifunctional coupling reagents, for example.

Another subject of the invention relates to compositions comprising a vector as defined above.

Advantageously, the vectors according to the invention may also be combined with one or more agents known for transfecting DNA. There may be mentioned by way of example the cationic lipids which possess advantageous properties. These vectors actually consist of a cationic polar part which interacts with the DNA, and a hydrophobic lipid part which promotes cellular penetration and renders the ionic interaction with DNA insensitive to the external medium. Specific examples

of cationic lipids are in particular the monocationic lipids (DOTMA: Lipofectin7), some cationic detergents (DDAB), lipopolyamines and in particular dioctadecylamidoglycyl spermine (DOGS) or 5-carboxyspermylamide of palmitoylphosphatidyl-ethanolamine (DPPES), whose preparation has been described, for example, in patent application EP 394 111. Another advantageous lipopolyamine family is represented by the compounds described in patent application WO 97/18185 incorporated into the present by way of reference. Numerous other cationic lipids have been developed and may be used with the vectors according to the invention.

Among the synthetic transfection agents developed, the cationic polymers of the polylysine and the DEAE-dextran type are also advantageous. It is also possible to use the polyethylenimine (PEI) and polypropylenimine (PPI) polymers which are commercially available and may be prepared according to the method described in patent application WO 96/02655.

In general, any synthetic agent known to transfect nucleic acid may be combined with the vectors according to the invention.

The compositions may, in addition, comprise adjuvants capable of combining with the vector according to the invention/transfection agent complexes

and of improving the transfecting power thereof. In another embodiment, the present invention therefore relates to compositions comprising a vector as defined above, one or more transfection agents as defined above and one or more adjuvants capable of combining with the vector according to the invention/transfection agent(s) complexes and of improving the transfecting power thereof. The presence of this type of adjuvant (lipids, peptides or proteins, for example) can advantageously make it possible to increase the transfecting power of the compounds.

In this regard, the compositions of the invention may comprise, as adjuvant, one or more neutral lipids whose use is particularly advantageous. The applicant has indeed shown that the addition of a neutral lipid makes it possible to enhance the formation of nucleolipid particles and to promote the penetration of the particle into the cell by destabilizing its membrane.

More preferably, the neutral lipids used within the framework of the present invention are lipids containing 2 fatty chains. In a particularly advantageous manner, natural or synthetic lipids which are zwitterionic or lacking ionic charge under physiological conditions are used. They may be chosen more particularly from dioleoylphosphatidylethanolamine

(DOPE), oleoylpalmitoylphosphatidylethanolamine (POPE), di-stearoyl, -palmitoyl, -mirystoyl phosphatidylethanolamines as well as their derivatives which are N-methylated 1 to 3 times, phosphatidyl-glycerols, diacylglycerols, glycosyldiacylglycerols, cerebrosides (such as in particular galacto-cerebrosides), sphingolipids (such as in particular sphingomyelins) or asialogangliosides (such as in particular asialoGM1 and GM2).

These different lipids may be obtained either by synthesis or by extraction from organs (for example the brain) or from eggs, by conventional techniques well known to persons skilled in the art. In particular, the extraction of the neutral lipids may be carried out by means of organic solvents (see also Lehninger, Biochemistry).

The applicant has demonstrated that it was also particularly advantageous to use, as adjuvant, a compound directly involved or otherwise in the condensation of the DNA (WO 96/25508). The presence of such a compound in the composition according to the invention makes it possible to reduce the quantity of transfecting compound, with the beneficial consequences resulting therefrom from the toxicological point of view, without any damaging effect on the transfecting activity. Compound involved in the condensation of the

nucleic acid is intended to define a compound which compacts, directly or otherwise, the DNA. More precisely, this compound may either act directly at the level of the DNA to be transfected, or may be involved at the level of an additional compound which is directly involved in the condensation of this nucleic acid. Preferably, it acts directly at the level of the DNA. In particular, this agent which is involved in the condensation of the DNA may be any polycation, for example polylysine. According to a preferred embodiment, this agent may also be any compound which is derived, as a whole or in part, from a histone, a nucleolin, a protamine and/or from one of their derivatives, as a whole or in part, of peptide units (KTPKKAKKP SEQ ID No.16) and/or (ATPAKKAA SEQ ID No. 17), it being possible for the number of units to vary between 2 and 10. In the structure of the compound according to the invention, these units may be repeated continuously or otherwise. They may thus be separated by linkages of a biochemical nature, for example one or more amino acids, or of a chemical nature.

The subject of the invention is also the use of the vectors as defined above for manufacturing a medicament intended to treat diseases by transfection of DNA into primary cells or into established cell lines. They may be fibroblast cells, muscle cells,

nerve cells (neurons, astrocytes, glial cells), hepatic cells, haematopoietic cell line (lymphocytes, CD34, dendritic cells, and the like), epithelial cells and the like, in differentiated or pluripotent form (precursors).

The vectors of the invention may, by way of illustration, be used for the *in vitro, ex vivo* or *in vivo* transfection of DNA encoding proteins or polypeptides.

For usage *in vivo*, either in therapy or for studying the regulation of the genes or the creation of animal models of pathological conditions, the compositions according to the invention can be formulated for administration by the topical, cutaneous, oral, rectal, vaginal, parenteral, intranasal, intravenous, intramuscular, subcutaneous, intraocular, transdermal, intratracheal or intraperitoneal route, and the like. Preferably, the compositions of the invention contain a vehicle which is pharmaceutically acceptable for an injectable formulation, in particular a direct injection into the desired organ, or for administration by the topical route (on the skin and/or the mucous membrane). They may be in particular isotonic sterile solutions, or dry, in particular freeze-dried, compositions which, upon addition, depending on the case, of sterilized water or of physiological saline,

allow the constitution of injectable solutions. The nucleic acid doses used for the injection as well as the number of administrations may be adapted according to various parameters, in particular according to the mode of administration used, the relevant pathological condition, the gene to be expressed, or the desired duration of treatment. As regards more particularly the mode of administration, it may be either a direct injection into the tissues or the circulatory system, or a treatment of cells in culture followed by their reimplantation *in vivo* by injection or transplantation.

The invention relates, in addition, to a method of transfection of DNA into cells comprising the following steps:

(1) synthesis of the oligonucleotide-targeting signal chimera according to the method described above,

(2) bringing the chimera synthesized in (1) into contact with a double-stranded DNA so as to form triple helices,

(3) optionally, complexing the vector obtained in (2) with one or more transfection agents and/or one more adjuvants, and

(4) bringing the cells into contact with the complex formed in (2) or, if applicable, in (3).

The cells may be brought into contact with the complex by incubating the cells with the said

complex (for uses *in vitro* or *ex vivo*), or by injecting the complex into an organism (for uses *in vivo*).

The present invention thus provides a particularly advantageous method for the treatment of diseases by administration of a vector according to the invention containing a nucleic acid capable of correcting the said disease. More particularly, this method is applicable to diseases resulting from a deficiency in a protein or peptide product, the administered DNA encoding the said protein or peptide product. The present invention extends to any use of a vector according to the invention for the *in vivo, ex vivo* or *in vitro* transfection of cells.

The invention also relates to any recombinant cell containing a vector as defined above. It involves preferably eukaryotic cells, for example yeast, animal cells, and the like. These cells are obtained by any technique allowing the introduction of a DNA into a given cell known to persons skilled in the art.

The following examples which are intended to illustrate the invention without limiting its scope make it possible to demonstrate other characteristics and advantages of the present invention.

FIGURES

Figure 1: Coupling of an oligonucleotide (ODN) and the

peptide maleimide-NLS.

Figure 2: Analysis, on a 15% polyacrylamide gel, of the oligonucleotide-peptide (Pso-GA_{19}-NLS) chimera by proteolytic action of trypsin.

1 = oligo Pso-GA_{19}-SH

2 = oligonucleotide-peptide Pso-GA_{19}-NLS chimera

3 = oligonucleotide-peptide Pso-GA_{19}-NLS chimera after digestion with trypsin

Figure 3: schematic representation of the plasmid pXL2813.

Figure 4: schematic representation of the plasmid pXL2652.

Figure 5: analysis on a 15% polyacrylamide gel of the formation of triple helices between the plasmid pXL2813 and the oligonucleotide-peptide Pso-GA_{19}-NLS chimera. The oligonucleotide pso-GA_{19}-NLS and the plasmid are mixed in a buffer containing 100 mM $MgCl_2$. The molar excess of oligonucleotide relative to the plasmid varies from 0 to 200. The mixture is photoactivated, after leaving overnight at 37EC, and then digested with two restriction enzymes which cut the plasmid on either side of the region of formation of the triple helices.

1 = no oligonucleotide

2 = molar excess of oligonucleotide relative to the plasmid of 15

3 = molar excess of oligonucleotide relative to the

plasmid of 50

4 = molar excess of oligonucleotide relative to the plasmid of 100

5 = molar excess of oligonucleotide relative to the plasmid of 200

6 = photoactivated oligonucleotide alone

7 = nonphotoactivated oligonucleotide alone

8 = molar excess of oligonucleotide relative to the plasmid of 30, the mixture not having been photoactivated.

Figure 6: Expression *in vitro* of the transgene (β-galactosidase) for tests carried out with the plasmid pXL2813 alone, the vector pXL2813-Pso-GA_{19}-NLS, and without plasmid.

Figure 7: Characterization of the peptide part of the oligonucleotide-peptide Pso-GA_{19}-NLS chimera by interaction with importin 60-GST (analysis on a 15% polyacrylamide gel).

1 = oligo Pso-GA_{19} (1μg)

2 = oligonucleotide-peptide Pso-GA_{19}-NLS chimera (1μg)

3 = supernatant recovered after incubation of glutathione-beads coated with importins 60 and the oligonucleotide Pso-GA_{19}, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant

4 = pellet recovered after incubation of glutathione-

beads coated with importins 60 and the oligonucleotide Pso-GA_{19}, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant

5 = supernatant recovered after incubation of the glutathione-beads coated with importins 60 and the oligonucleotide-peptide Pso-GA_{19}-NLS chimera, and separation of the pellet of beads (containing the components which react with the importins) from the supernatant

6 = pellet recovered after incubation of the glutathione-beads coated with importins 60 and the oligonucleotide-peptide Pso-GA_{19}-NLS chimera, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant.

Figure 8: Analysis on a 15% polyacrylamide gel of the oligonucleotide-peptide (GA_{19}-NLS) chimera by the proteolytic action of trypsin.

1 = oligo GA_{19}-SH (200 ng)

2 = oligonucleotide-peptide GA_{19}-NLS chimera (1μg) before purification by high-performance liquid chromatography

3 = oligonucleotide-peptide GA_{19}-NLS chimera (1μg) after purification

4 = oligonucleotide-peptide GA_{19}-NLS chimera after

digestion with trypsin (1μg).

Figure 9: Graphical representation of the kinetics of formation of the triple helices (% of triple helix sites occupied as a function of time) between the plasmid pXL2813 and the chimera GA_{19}-NLS.

Figure 10: Characterization of the peptide part of the oligonucleotide-peptide GA_{19}-NLS chimera by interaction with the importin 60-GST.

1 = oligonucleotide-peptide GA_{19}-NLS chimera,

2 = oligo GA_{19},

3 = supernatant recovered after incubation of the glutathione-beads coated with importins 60 and the oligonucleotide-peptide GA_{19}-NLS chimera, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant,

4 = pellet recovered after incubation of the glutathione-beads coated with importins 60 and the oligonucleotide-peptide GA_{19}-NLS chimera, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant,

5 = supernatant recovered after incubation of the glutathione-beads coated with importins 60 and the oligonucleotide-peptide GA_{19} and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant,

6 = pellet recovered after incubation of the

SUB A7

glutathione-beads coated with importins 60 and the oligonucleotide GA_{19}, and separation of the pellet of beads (containing the components which interact with the importins) from the supernatant.

Figure 11: Schematic representation of the plasmid pXL2997.

Figure 12: Graphical representation of the kinetics of formation of the triple helices (% of triple helix sites occupied as a function of time) between the plasmid pXL2997 and the chimera pim-NLS.

Figure 13: Histogram representing the β-galactosidase activity *in vivo* in human pulmonary tumours H1299 of the plasmids pXL2813 (indicated Bgal in the figure) and pXL2813-Pso-GA_{19}-NLS (indicated NLS-Bgal in the figure) in RLU ("Relative Light Unit") per tumour.
The transfection was performed using the electrotransfer techniques as described in applications WO 99/01157 and WO 99/01158.

MATERIALS AND METHODS

1. Coupling of the oligonucleotides and the peptides

Oligonucleotides

The oligonucleotides used are the sequence 5'-AAGGAGAGGAGGGAGGGAA-3' (SEQ ID No. 4) 19 bases long and referenced under the name "GA_{19}" in the text which

follows, or the sequence 5'-GGGGAGGGGGAGG-3' (SEQ ID No. 15) 13 bases long and referenced under the name "pim" in the text which follows (because it is the sequence of the protooncogene pim-1).

The oligonucleotides noted GA_{19}-SH or pim-SH have the same sequences as GA_{19} and pim, respectively, and a thiol group at the 5' end, with a spacer of six carbon between the thiol and the phosphate of the 5' end. The oligonucleotides noted Pso-GA_{19}-SH have a thiol group at the 3' (SH) end, and in addition a psoralen at the 5' (Pso) end, with a spacer of six carbons between the psoralen and the phosphate of the 5' end. The oligonucleotides noted Pso-GA_{19} do not have a thiol group.

The nomenclature used is summarized in Table I below:

name of the oligonucleotide	modification of the 3' end	modification of the 5' end
GA_{19}	none	none
pim	none	none
GA_{19}-SH	none	thiol group
pim-SH	none	thiol group
Psi-GA_{19}	none	psoralen
Pso-GA_{19}-SH	thiol group	psoralen

Table 1

These freeze-dried oligonucleotides are taken up in a 100 mM trimethylammonium acetate buffer, pH = 7.

Peptides

The peptides used for the couplings are synthesized by an automatic machine in the solid phase. They contain:

- either the nuclear localization sequence of the SV40 T antigen (PKKKRKV, SEQ ID No. 11),
- or the same sequence, mutated (PK**N**KRKV, SEQ ID No. 14) which allows neither targeting nor nuclear transport (because of the mutation).

These peptides also carry a spacer of four amino acids at the N-terminal end: KGAG. The N-terminal lysine is chemically modified: it contains a maleimide group and the ε carbon and a protecting group 9-fluorenylmethyloxycarbonyl (Fmoc) on the amine of the α carbon. This Fmoc group absorbs at 260 nm, which makes it possible to monitor the peptide by reversed-phase high-performance liquid chromatography. The C-terminal group is also protected ($CONH_2$ group), the protection being added at the end of peptide synthesis.

The representation of the peptides is indicated in Table II below:

SUB A10

name of the peptide	sequence and modifications
maleimide-NLS	maleimide-KGAGPKKKRKV-$CONH_2$] Fmoc
maleimide-NLSmutated	maleimide-KGAGPKNKRKV-$CONH_2$] Fmoc

Table II

The freeze-dried peptides are picked up in a 100 mM triethylammonium acetate buffer, pH = 7. The concentration is 0.4 mg/ml.

Couplings

For the coupling with the oligonucleotides, the strategy used consists in reacting the thiol group carried by the oligonucleotide and the maleimide group carried by the peptide [Eritja, R., et al., *Synthesis of defined peptide-oligonucleotide hybrids containing a nuclear transport signal sequence*, Tetrahedron, 1991, **47**(24), pp. 4113-4120].

The oligonucleotide is added to the peptide solution in an equimolar quantity and the reaction medium is left for 2 hours at room temperature. The oligonucleotide-peptide chimera is purified by reversed-phase high-pressure liquid chromatography on a Vydac C8 column containing spheroidal silica having a diameter of 5 μm and a porosity of 300 Å. Use is made

of a 0.1 M triethylammonium acetate (TEAA) buffer and an acetonitrile gradient passing from 5% to 50% over 35 minutes. The products are detected at 260 nm.

Analysis of the chimeras by digestion with trypsin

The oligonucleotide-peptide conjugates are subjected to the proteolytic action of trypsin which makes it possible to reveal the peptide part of the chimera [Reed, M.W. et al., *Synthesis and evaluation of nuclear targeting peptide-anti-sense oligodeoxynucleotide conjugates*, Bioconjugate Chemistry, 1995, 6, pp.101-108]. Solutions containing 1 μg of oligonucleotide-peptide in 7 μl of 0.1 M triethylammonium (TEAA) purification buffer are mixed with 1 μl of a trypsin solution (5 mg/ml). 1 μl of 100 mM Tris-HCl buffer, pH = 9, and 1 μl of 500 mM EDTA are added thereto. After digesting for one hour, the samples are deposited into the wells of a 15% polyacrylamide, 7 M urea, gel. The electrophoresis is performed in 100 mM Tris buffer, pH 8.3 containing 90 mM boric acid and 1 mM EDTA. The nucleic acids are visualized by silver staining using a Biorad kit.

2. Formation of triple helices with the chimeras

Plasmid

The plasmid used to study the formation of triple helices with the GA_{19}-peptide chimeras is called pXL2813 (7257 pb, cf Figure 3). This plasmid expresses the β-galactosidase gene under the control of the strong promoter of the cytomegalovirus (CMV) early genes, as well as the ampicillin resistance gene.

Upstream of the promoter, between positions 7238 and 7256, the GA_{19} sequence (5'-AAGGAGAGGAGGGAGGGAA-3', SEQ ID No. 4) was cloned according to conventional molecular biological techniques.

It was cloned into the plasmid pXL2652 (of 7391 bp and which is schematically represented in Figure 4) which expresses the β-galactosidase gene under the control of the strong promoter of the cytomegalovirus (CMV) early genes, as well as the ampicillin resistance gene. This promoter is derived from the pCDNA3, the LacZ and its polyA are derived from the pCH110, and the remainder is derived from pGL2.

The sequences were cloned upstream of the promoter between the unique cleavage sites for the enzymes MunI and XmaI. For that, the two complementary oligonucleotides containing the sequence to be cloned

6651 (5'-AATTGATTCCTCTCCTCCCTCCCTTAC-3'SEQ ID No.18) and 6652 (3'-CTAAGGAGAGGAGGGAGGGAATGGG-5' SEQ ID No.19) were heated for 5 minutes at 95EC and then hybridized by allowing the temperature to fall slowly. The plasmid pXL2652 was then digested with the enzymes MunI and XmaI for 2 hours at 37EC and the products of this double digestion were separated by electrophoresis on a 1% agarose gel followed by staining with ethidium bromide.

The fragment of interest for the remainder of the cloning was eluted according to the Jetsorb protocol (Genomed) and 200 ng of this fragment were linked to 10 ng of the mixture of oligonucleotides hybridized by T4 ligase, for 16 hours at 16EC.

Competent E. coli bacteria of the DH5α strain were transformed by eletroporation with the reaction product, plated on Petri dishes containing LB medium and ampicillin. The ampicillin-resistant clones were selected and the DNA was extracted by alkaline lysis and analysed on a 1% agarose gel. A clone corresponding in size to the expected product was sequenced.

In the case where the oligonucleotide is pim, the plasmid used to study priple helix formation is pXL2997 represented in Figure 11. This plasmid expresses the gene for β-galactosidase under the control of the strong promoter of the Cytomegalovirus

(CMV) early genes, as well as the gene for resistance to Ampicillin.

Upstream of the promoter, the sequence pim (SEQ ID No. 15) was cloned according to conventional molecular biology methods.

Formation of the triple helices

The formation of the triple helices is performed by mixing the plasmid pXL2813 (3 pmol of plasmid, that is to say 15 μg) or pXL2997, depending on the case, and variable quantities of oligonucleotides or of chimeras in a 100 mM Tris-HCl buffer, pH = 7.5, containing 100 mM $MgCl_2$.

Photoactivation of the triple helices

After incubating overnight, the mixture is irradiated, in ice, for 15 minutes, using a monochromatic lamp having a wavelength of 365 nm (Biorad). The product of photoactivation is digested with the restriction enzymes MfeI and SpeI and analysed on a 15% polyacrylamide, 7 M urea, gel with 100 mM Tris migration buffer, pH = 8.3, containing 90 mM boric acid and 1 mM EDTA. The nucleic acids are visualized by silver staining using a Biorad kit [Musso, M., J.C. Wang, and M.W.V. Dyke, *In vivo persistence of DNA triple helices containing psoralen-conjugated*

oligodeoxyribonucleotides, Nucleic Acid Research, 1996, 24(24), pp. 4924-4932].

Method of studying the triple helices

This method is based on the principle of extrusion chromatography of solutions containing the plasmid and the oligonucleotide which are capable of forming a triple helix. The exclusion columns used (Columns, Linkers 6, Boehringer Mannheim) are composed of Sepharose beads and have an exclusion limit of 194 base pairs, which makes it possible to retain in the column the oligonucleotides not paired with the plasmids.

In a first instance, the oligonucleotides are radioactively labelled at their 3' end with Terminal Transferase with the aid of [α-^{35}S]dATP. The protocol used is from Amersham: 10 pmol of oligonucleotides are incubated for 2 hours at 37EC, with 50 μCi of [α-^{35}S]dATP in the presence of 10 units of Terminal Transferase, in a volume of 50 μl of buffer containing sodium cacodylate. The percentage of labelled oligonucleotides is evaluated according to the following method: 1 μl of a sample, diluted 1/100, of the solution after labelling is deposited on Whatman DE81 paper, in duplicate. One of the two papers is washed twice for 5 minutes with 2HSSC, for 30 seconds

with water and for 2 minutes with ethanol. The radioactivities of the two papers are compared. The radioactivity of the washed paper corresponds to the ^{35}S effectively incorporated.

The formation of the triple helices occurs in a volume of 35 µl. The concentrations of plasmids (40 nM) and of oligonucleotides (20 nM), the buffer used (100 mM Tris-HCl pH = 7.5, 50 mM of $MgCl_2$) and the temperature (37EC) are fixed whereas the incubation time varies from 1 to 24 hours. The Sepharose columns are equilibrated, before use, with the reaction buffer and centrifuged at 2200 rpm for 4 minutes, in order to cause them to settle. 25 µl of the reaction medium are deposited on the columns and the latter are centrifuged under the same conditions as above. 25 µl of the reaction buffer are then deposited on the columns which are again centrifuged. The eluate is recovered.

The radioactivity contained in 5 µl of the reaction medium, noted cpm(deposit), and that contained in the eluate, noted cpm(eluate), are evaluated, which makes it possible to estimate the percentage of oligonucleotides eluted:

% of oligonucleotides eluted = cpm(eluate)/ [5 H cpm(deposit)] H 100.

The percentage of plasmids which are effectively eluted during the experiments is evaluated

by estimating the optical density at 260 nm of the eluate and that of the deposit, which makes it possible to calculate the percentage of oligonucleotides attached to the total number of plasmids.

% of oligonucleotides attached = [% oligonucleotides eluted/% plasmids eluted] H 100.

This parameter makes it possible to evaluate the percentage of sites of formation of triple helices (there is one per plasmid pXL2813 or pXL2997) which are effectively occupied, taking into account the concentrations of plasmids (noted [plasmid]) and of oligonucleotides (noted [oligo]) used during the reaction of formation of the triple helices:

% triple helix sites occupied = % of oligonucleotides attached H [oligo]/[plasmid].

3. Interaction with the importins

Recombinant proteins

The importin 60 subunit used to study the interaction with the oligonucleotide-peptide (NLS or NLS*mutated*) conjugates is of murine origin and is fused with Glutathione S-Transferase (GST). The sequence of importin 60 was cloned into a vector pGEX-2T in order to fuse it with GST. The recombinant protein was produced in Escherichia coli [Imamoto, N. et al., *In*

vivo evidence for involvement of a 58kDa component of nuclearpore-targeting complex in nuclear protein import, The EMBO Journal, 1995, 14(15), pp. 3617-3626].

Interactions with the recombinant proteins

All the interaction experiments are carried out in the binding buffer (20 mM HEPES, pH = 6.8, 150 mM potassium acetate, 2 mM magenesium acetate, 2 mM DTT and 100 μg/ml BSA)

In a first instance, the recombinant proteins are incubated in the presence of Sepharose beads coated with glutathion (Pharmacia Biotech), 1 μg of recombinant protein is used for 10 μl of beads. After incubating for 30 minutes at room temperature in 500 μl of binding buffer, the mixture is centrifuged at 2000 g for 30 seconds, and the supernatant is removed. The beads are washed five times by resuspending in 500 μl of binding buffer followed by centrifugation, as described above. The beads are resuspended in binding buffer in order to obtain a suspension containing 50% of beads coated with recombinant proteins.

In a second instance, 60 μl of the suspension containing 50% of beads coated with recombinant proteins are incubated with 2 μg of oligonucleotide or of oligonucleotide-peptide, in a volume of 500 μl of binding buffer. After incubating for 30 minutes at room

temperature, the mixture is centrifuged at 2000 g for 30 seconds, and the supernatant is removed. 30 µl of the supernatant are collected in order to analyse the fraction not attached to the beads. The beads are washed five times by resuspending in 500 µl of binding buffer followed by centrifugation, as described above. The beads are resuspended in 15 µl of loading buffer (0.05% bromophenol blue, 40% sucrose, 0.1 M EDTA, pH = 8, 0.5% sodium lauryl sulphate) and heated for 10 minutes at 90EC. The content of the supernatant and of the pellet is analysed on a 15% polyacrylamide, 7 M urea, gel, with a 100 mM Tris migration buffer, pH 8.3, containing 90 mM boric acid and 1 mM EDTA. The nucleic acids are visualized by silver staining using a Biorad kit [Rexach, M. and G. Blobel, *Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins*, Cell, 1995, 83, pp. 683-692].

4. Transfection of cells

Cell culture

The cell type used is NIH 3T3 (ATCC CRL-1658). It consists of mouse fibroblasts. These cells are cultured in a modified Dulbecco's medium, with 4.5 g/l glucose (DMEM - Gibco), 2 mM glutamine, penicillin (100 U/ml) and streptomycin (100 µg/ml), and

10% foetal calf serum (Gibco). They are incubated at 37EC in an incubator containing 5% CO_2.

Transfection

One day before the transfection, the wells of a 24-well plate are inoculated with 50,000 cells per well. The vectors are diluted in 150 mM NaCl and mixed with a cationic lipid (the compound having the condensed formula $H_2N(CH_2)_3NH(CH_2)_4NH(CH_2)_3NHCH_2COGlyN[(CH_2)_{17}CH_3]_2$ described in patent application WO 97/18185 under the number (6)) diluted in 150 mM NaCl. The mixture is made with 6 nmol of lipid per microgram of plasmid. This mixture is diluted 1/10 in serum-free culture medium and deposited on the cells. After incubating at 37EC in an incubator containing 5% CO_2 for two hours, 10% foetal calf serum is added.

Quantification of the β-galactosidase

After incubating for 48 hours, the cells are washed twice with PBS and lysed with 250 µl of lysis buffer (Promega). The β-galactosidase is quantified according to the protocol "Lumigal β-Galactosidase genetic reporter system" (Clontech). The activity is measured on a Lumat LB9501 luminometer (Berthold). The quantity of proteins is measured with the BCA kit

(Pierce).

Examples

This example illustrates the possibility of coupling the oligonucleotide Pso-GA_{19}-SH to the peptide maleimide-NLS.

The oligonucleotide Pso-GA_{19}-SH, of sequence 5'-AAGGAGAGGAGGGAGGGAA-3', SEQ ID No.4) with a thiol group at the 3' end, was coupled to the peptide NLS which carries a maleimide group at its N-terminal end according to the method described above in the "Materials and Methods" part under the section "coupling of the oligonucleotides and the peptides".

The coupling is monitored by reversed-phase high-performance liquid chromatography (HPLC). It occurs with an equimolar stoechiometry: in two hours, the coupling is complete and the chimera is purified by HPLC.

The reaction sheme for the coupling is represented in Figure 1.

The oligonucleotide-peptide (Pso-GA_{19}-NLS) chimera was analysed by polyacrylamide gel electrophoresis after proteolytic action of trypsin which makes it possible to reveal the presence of the peptide part of the chimera, after migration on a denaturing polyacrylamide gel and silver staining of

the nucleic acids (as indicated in the "Materials and Methods" part under the section "Analysis of the chimeras by digestion with trypsin").

The chimera Pso-GA_{19}-NLS exhibits a retarded electrophoretic migration compared with the oligonucleotide Pso-GA_{19}-SH, and the product of proteolytic digestion is visualized at an intermediate level between the levels of migration of Pso-GA_{19}-NLS and of Pso-GA_{19}-SH, as shown in Figure 2.

The chimera Pso-GA_{19}-NLS therefore contains a peptide part which is accessible to trypsin. These results clearly show that the coupling between the oligonucleotide and the peptide occurred, and the coupling yield is high.

Example 2

This example illustrates the formation of triple helices between the plasmid pXL2813 and the chimera Pso-GA_{19}-NLS which is modified by a photoactivable alkylating agent. This example also indicates the proportion of plasmids modified according to the molar excess of oligonucleotides relative to the plasmid.

The plasmid pXL2813, represented in Figure 3, comprises the homopurine sequence complementary to GA_{19} which is capable of forming a triple helix with the

oligonucleotides GA_{19}, Pso-GA_{19}, Pso-GA_{19}-SH or Pso-GA_{19}-NLS. Divalent cations such as Mg^{2+} stabilize these triple helices. The oligonucleotide Pso-GA_{19}-NLS and the plasmid are mixed in a buffer containing 100 mM of $MgCl_2$. The molar excess of oligonucleotide relative to the plasmid varies from 0 to 200.

SUB A15

After incubating for 12 hours at 37EC, the mixture is photoactivated for 15 minutes (as indicated in the "Materials and Methods" part under the section "Photoactivation of the triple helices"), and then digested with the two restriction enzymes MfeI and SpeI which cut the plasmid on either side of the region of formation of the triple helices. A nucleic acid fragment is thus released which contains 70 base pairs after digestion of the nonmodified plasmid pXL2813. The fragment obtained with the plasmid and the oligonucleotide forming a triple helix covalently bound to the double helix has 70 base pairs plus 19 bases of the oligonucleotide. By migration on a denaturing polyacrylamide gel, it is thus possible to separate these fragments of different size: the fragments derived from the plasmids modified by a triple helix have a shorter migration distance than the fragments derived from nonmodified plasmids. It is therefore possible, by electrophoresis on a denaturing polyacrylamide gel, to quantify the proportion of

modified plasmids according to the molar excess of oligonucleotide relative to the plasmid.

The results are indicated in Figure 5. For a molar excess of oligonucleotide relative to the plasmid greater than 50, all the plasmids are modified and are combined with an oligonucleotide Pso-GA_{19}-NLS. Without photoactivation, the retardation of the digestion fragment is lost in a denaturing gel.

It thus appears that the triple helices formed with the oligonucleotides Pso-GA_{19}-SH or Pso-GA_{19}-NLS are therefore indeed covalently bound to the double helix after photoactivation.

Moreover, by digesting the remainder of the plasmid backbone and by analysing as above, it is possible to check that the photoactivation does not result in a nonspecific covalent binding of the oligonucleotide Pso-GA_{19}-NLS outside the region containing the sequence capable of forming a triple helix.

These results clearly demonstrate the conditions which make it possible to specifically and covalently combine a peptide sequence with a plasmid, and is outside the promoter regulating the expression of the transgene, which prevents the expression of the transgene from being affected.

Example 3

This example illustrates the capacity of the transgene to be expressed *in vitro*, although the plasmid is modified.

The expression of β-galactosidase by the plasmids pXL2813 which are nonmodified or which are combined with an oligonucleotide Pso-GA_{19}-NLS was thus compared by transfection of NIH3T3 cells.

The results are reported in Figure 6, and the mean values obtained (∀ standard deviation) are assembled in the following Table III:

plasmid	expression of the transgene (in RLU/μg of protein)
pXL2813	253759 ∀ 48545
pXL2813-Pso-GA_{19}-NLS	473984 ∀ 111476
without plasmid	129 ∀ 85

Table III

It can be observed that the expression of the transgene increases following the modification provided by the covalent attachment of the triple helix upstream of the promoter.

These results clearly demonstrate that the formation of the triple helices does not affect the

expression of the transgene *in vitro*. In contrast, by virtue of the nuclear homing of the plasmid because of its coupling with the targeting sequence NLS, more plasmids reached the nucleus, resulting in an increase in the transfection efficiency.

Example 4

This example is intended to verify that the *in vivo* expression is not inhibited by the presence of a targeting signal associated with the plasmid via a triple helix.

The experiment consists in transfecting pXL2813 and pXL2813-Pso-GA_{19}-NLS into human pulmonary tumours H1299 using the electrotransfer techniques described in applications WO 99/01157 and WO 99/01158. The experiment was carried out in 18 to 20 g female nude mice.

The mice are monolaterally implanted with 20 mm^3 H1299 tumour grafts. The tumours develop, reaching a volume of 200 to 300 mm^3. The mice are separated according to the size of their tumours and divided into homogeneous groups. The mice are then anaesthetized with a Ketamine/Xylazine mixture (commercially available). The plasmid solution (8 μg of DNA/40 μl of 150 mM sodium chloride) is longitudinally injected at the periphery of the tumour with the aid of

a Hamilton syringe.

The lateral faces of the tumour are covered with conducting gel and the tumour is placed between two stainless steel flat electrodes 0.45 to 0.7 cm apart. The electrical pulses are applied with the aid of a commercial square pulse generator (Electro-pulsateur PS 15, Jouan, France) 20 to 30 seconds after the injection. An oscilloscope makes it possible to control the intensity in volts, the duration in milliseconds and the frequency in hertz of the pulses which should be delivered at 500 V/cm for 20 milliseconds at 1 hertz.

For the evaluation of the tumour transfection, 10 and 7 mice respectively for the plasmid pXL2813 and pXL2818-Pso-GA_{19}-NLS are humanely killed 2 days after the injection of the plasmid. The tumours are collected, weighed and ground in a lysis buffer (Promega) supplemented with antiproteases (Complete, Boehringer). The suspension obtained is centrifuged in order to obtain a clear supernatant. After having incubated 10 µl of this supernatant with 250 µl of reaction buffer (Clontech) for 1 hour in the dark, the β-galactosidase is measured with the aid of a commercial luminometer. The results are expressed in total RLU ("Relative Light Unit") per tumour.

It is observed that the plasmid pXL2813-Pso-

GA_{19}-NLS expresses the transgene *in vivo* at a level greater than or equal to that obtained with the unmodified plasmid pXL2813 (see Figure 13).

Example 5

This example illustrates the characterization of the peptide part of the Pso-GA_{19}-NLS conjugates, that is to say the verification of the targeting properties of the NLS peptide combined with the constructs according to the invention.

The peptide sequence used, the NLS signal of the SV40 T antigen, is recognized by receptors of the α-karyopherin family. The murine equivalent, called importin 60, fused with a glutothione S-transferase group, was used to characterize the oligonucleotide-peptide conjugates. It was performed according to the method described in the "Materials and Methods" part under the section "Interaction with the importins".

The interactions between the glutathion-beads coated with importins 60 and the oligonucleotides GA_{19}-NLS or Pso-GA_{19}-NLS were studied. After incubating for 30 minutes at room temperature, the pellet of beads (containing the components which interact with the importins) is separated from the supernatant.

The result of this characterization is reported in Figure 7. This figure indicates that the oligonucleotide Pso-GA_{19}-NLS is combined with the

glutathion beads, whereas the oligonucleotide Pso-GA_{19} remains in the supernatant.

This clearly shows the capacity of the oligonucleotide Pso-GA_{19}-NLS to interact with the α-importin. The peptide part of the oligonucleotide-peptide chimeras is therefore indeed recognized by its receptor, which means that the peptide effectively fulfils its role of targeting signal.

Example 6

This example illustrates the possibility of coupling the oligonucleotide GA_{19}-SH to the peptide maleimide-NLS. Unlike Example 1, the chimera is not modified by a photoactivable alkylating agent.

The oligonucleotide GA_{19}-SH, of sequence 5'-AAGGAGAGGAGGGAGGGAA-3' (SEQ ID No. 4), with a thiol group at the 5' end, was coupled to the peptide maleimide-NLS which possesses a maleimide group at its N-terminal end under the same conditions as for the oligonucleotide Pso-GA_{19}-SH (see Example 1).

The coupling is monitored by reversed-phase high-performance liquid chromatography (HPLC). It occurs with an equimolar stoichiometry: in two hours, the coupling is complete. The chimera is then purified by HPLC.

The oligonucleotide-peptide chimera was

analysed by proteolytic action of trypsin as described in the "Materials and Methods" part.

The result is represented in Figure 8. It is comparable to that obtained with the oligonucleotide Pso-GA_{19}-SH.

Example 7

This example illustrates the possibility of forming triple helices between the plasmid pXL2813 and the $GA-_{19}$-NLS chimera in the absence of alkylating agent.

A kinetics of formation of triple helices between the plasmid pXK2813 and the $GA-_{19}$-NLS chimera, obtained as described in Example 5, was carried out and studied according to the technique described in the "Materials and Methods" part, under the section "Formation of triple helices with the chimeras".

Figure 9 represents the kinetics of formation of the triple helices.

It appears that under the conditions of a plasmid concentration of 40 nM and an oligonucleotide concentration of 20 nM, a stable triple helix is formed.

Example 8

This example illustrates the characterization

of the peptide part of the GA_{19}-NLS chimera.

The peptide sequence used, the NLS signal of the SV40 T antigen, is recognized by receptors of the α-karyopherin family, as already mentioned in Example 4. The murine equivalent, called importin 60, fused with a glutathione S-transferase group, was used to characterize the oligonucleotide-peptide conjugates. It was performed as described in "Materials and Methods", under the section "Interactions with the importins".

The interactions between the glutathione-beads coated with importins 60 and the oligonucleotides GA_{19} or GA_{-19}-NLS were studied. After incubating for 30 minutes at room temperature, the pellet of beads (containing the components which interact with the importins) is separated from the supernatant.

The results are indicated in Figure 10. It appears that the oligonucleotide GA_{19}-NLS is combined with the glutathione beads, whereas the oligonucleotide GA_{19} remains in the supernatant.

This clearly shows the capacity of the oligonucleotide GA_{-19}-NLS to interact with the α-importin. The peptide part of the oligonucleotide-peptide chimeras is therefore indeed recognized by its receptor, and fulfils, here again, its role as targeting signal.

Example 9

This example illustrates the possibility of coupling the oligonucleotide pim-SH to the peptide maleimide-NLS.

The oligonucleotide pim-SH, of sequence 5'-GGGGAGGGGGAGG-3' (SEQ ID No. 15), with a thiol group at the 5' end, was coupled to the peptide maleimide-NLS which possesses a maleimide group at its N-terminal end under the same conditions as for the oligonucleotide GA_{19}-SH (see Example 5).

The coupling is monitored by reversed-phase high-performance liquid chromatography (HPLC). It occurs with an equimolar stoichiometry: in two hours, the coupling is complete. The chimera is then purified by HPLC.

Example 10

This example illustrates the possibility of forming triple helices between the plasmid pXL2997 and the pim-NLS chimera in the absence of alkylating agent.

A kinetics of formation of triple helices between the plasmid pXL2997 and the pim-NLS chimera (formation obtained as described in Example 8), was carried out and studied according to the technique described in the "Materials and Methods" part, under the section "Formation of triple helices with the

chimera".

Figure 12 represents the kinetics of formation of the triple helices.

It appears that under the conditions of plasmid concentrations of 40 nM and of oligonucleotide concentrations of 20 nM, a stable triple helix is formed.

The invention having now been completely described, it is evident that numerous modifications can be made to the present invention without as a result parting from the spirit of the invention and from the scope of the claims which follow.